

# **SQL Server Oefeningen Performantie**

Subtitel

## Installation of test database Stack Overflow

All data of the famous forum stackoverflow is stored in a SQL Server database. We will use a fragment of this database (years 2008-2010). Because of the size we don't deliver a backup file but we restore the database directly from the datafile `StackOverflow2010.mdf`. Copy this file in your SQL Server data directory (something like `C:\Program Files\Microsoft SQL Server\MSSQL14.MSSQLSERVER\MSSQL\DATA`). Now, in SSMS, right mouse click on Databases, Attach, Add and browse to the mdf file.

### 1. Explore the data

Draw the database diagram. Guess, from the tables names, which table is the largest (in terms of number of records). How many records does this table have?

### 2. Exercises

1. Perform a `SELECT * FROM posts`. How many seconds does it take on your computer? Why does it take so long? How many seconds does `"SELECT id FROM posts"` take? Why is there such a difference?
2. Find the best way to select `id`, `body` and `lasteditordisplayname` for all posts created in 2008.
3. Order posts by score and (in case of equal score) `commentcount` both in descending order. Show `id`, `score`, `commentcount` and `title` in the most efficient way. How can you check if your result is executed in the most efficient way? Is the actual table used?
4. Create an index on `title`. Then explain the difference (in the execution plan) between following queries:

```
select id,title
from posts
where title like '%php%';
```

```
select id,title
from posts
where title like 'php%';
```

5. Optimize following query by creating an index. Check your execution plan before and after the index creation. Point out why the after-situation is far better.

```
select id,title,answercount
from posts
where answercount > 10
order by answercount desc;
```